REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
02-05-2013	Final Report	16-Sep-2008 - 15-Sep-2012

4. TITLE AND SUBTITLE
Statistical Relational Learning (SRL) as an Enabling Technology for Data Acquisition and Data Fusion in Video

5a. CONTRACT NUMBER
W911NF-08-1-0466

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Lise Getoor, David Jacobs

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Maryland - College Park
Research Administration
3112 Lee Building
College Park, MD 20742 -5141

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
54637-NS.11

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
While our ability to gather vast amounts of video data is growing at a staggering rate, our ability to effectively store, process, and analyze this video has not kept pace. It is therefore necessary to develop automatic methods for allocating limited resources in video understanding. It particular, it is important to reason about which portions of video require expensive analysis and storage. This project aims to make these inferences using new and existing tools from Statistical Relational Learning (SRL). SRL is a recently emerging technology that enables the effective

15. SUBJECT TERMS
statistical relational learning, data acquisition, active learning, collective classification

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		Lise Getoor
					19b. TELEPHONE NUMBER
					301-405-2691

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Statistical Relational Learning (SRL) as an Enabling Technology for Data Acquisition and Data Fusion in Video

ABSTRACT

While our ability to gather vast amounts of video data is growing at a staggering rate, our ability to effectively store, process, and analyze this video has not kept pace. It is therefore necessary to develop automatic methods for allocating limited resources in video understanding. It particular, it is important to reason about which portions of video require expensive analysis and storage. This project aims to make these inferences using new and existing tools from Statistical Relational Learning (SRL). SRL is a recently emerging technology that enables the effective integration of statistical or probabilistic information, with relational or logical domain information, providing the ability to reason collectively about large, complex, interacting domains.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received   Paper

TOTAL:

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received   Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

SRL Tutorials given:
Invited Tutorial, Conference on Neural Information Processing (NIPS), December 2012.
“Learning Statistical Models from Relational Data," ACM International Conference on Management of Data (SIGMOD), Athens, GR, June, 2011.
“Exploiting Statistical & Relational Information on the Web and in Social Media," Eleventh SIAM International Conference on Data Mining (SDM), Phoenix, AZ, April, 2011.
“Exploiting Statistical & Relational Information on the Web and in Social Media,” Fourth ACM International Conference on Web Search and Data Mining (WSDM), Hongkong, CH, February, 2011.
“Exploiting Statistical & Relational Information on the Web and in Social Media: Applications, Techniques, and New Frontiers," National Conference on Artificial Intelligence (AAAI), joint with Lily Mihalkova, Atlanta, GA, July, 2010.

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper
TOTAL:	

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received		Paper
05/01/2013	2.00	Daozheng Chen, Mustafa Bilgic, Lise Getoor, David Jacobs, Lilyana Mihalkova, Tom Yeh. Active Inference for Retrieval in Camera Networks, IEEE Workshop on Person Oriented Vision. 2011/01/07 00:00:00, . : ,
05/01/2013	1.00	Daozheng Chen, Mustafa Bilgic, Lise Getoor, David Jacobs. Efficient Resource-constrained Retrospective Analysis of Long Video Sequences, NIPS Workshop on Adaptive Sensing, Active Learning and Experimental Design: Theory, Methods and Applications. 2012/12/11 00:00:00, . : ,
05/01/2013	4.00	Daozheng Chen, Dhruv Batra, Bill Freeman, Micah K. Johnson.. Group Norm for Learning Latent Structural SVMs, NIPS Workshop on Optimization in Machine Learning. 2012/12/16 00:00:00, . : ,
05/01/2013	5.00	Ronen Basri, Pedro Felzenszwalb, Ross Girshick, David Jacobs, Caroline Klivans. Visibility Constraints on Features of 3D Objcts,’, 0IEEE Conference on Computer Vision and Pattern Recognition (CVPR),. 2009/06/20 00:00:00, . : ,
05/01/2013	8.00	Mustafa Bilgic, Lise Getoor. Effective Label Acquisition for Collective Classification, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2008/08/24 00:00:00, . : ,
05/01/2013	9.00	Mustafa Bilgic, Lilyana Mihalkova, Lise Getoor. Active Learning for Networked Data, International Conference on Machine Learning. 2010/06/25 00:00:00, . : ,
TOTAL:		**6**

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received		Paper
05/01/2013	3.00	Daozheng Chen, Mustafa Bilgic, Lise Getoor, David Jacobs. Dynamic Processing Allocation in Video, IEEE Transactions on Pattern Analysis and Machine Intelligence (11 2011)
05/02/2013	10.00	Mustafa Bilgic, Lise Getoor. Reflect and Correct: A Misclassification Prediction Approach to Active Inference, ACM Transactions on Knowledge Discovery from Data (11 2009)

TOTAL: **2**

Number of Manuscripts:

Books

Received Paper

TOTAL:

Patents Submitted

Patents Awarded

Awards

Best Paper Award, KDD 2008

Graduate Students

NAME	PERCENT_SUPPORTED	Discipline
Mustafa Bilgic	0.50	
Daozheng Chen	0.50	
Galileo Namata	0.02	
Carlos Castillo	0.06	
Anne Jorstad	0.02	
New Entry	0.00	
FTE Equivalent:	**1.10**	
Total Number:	**6**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Lise Getoor	0.05	
David Jacobs	0.06	No
FTE Equivalent:	**0.11**	
Total Number:	**2**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED	Discipline
Aditya Malik	0.00	Computer Science
FTE Equivalent:	**0.00**	
Total Number:	**1**	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 0.00

Names of Personnel receiving masters degrees

NAME	
Total Number:	

Names of personnel receiving PHDs

NAME	
Mustafa Bilgic	
Daozheng Chen	
Galileo Namata	
Total Number:	**3**

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

See attachment

Technology Transfer

Statistical Relational Learning (SRL) as an Enabling Technology for Data Acquisition and Data Fusion in Video
Proposal Number W911NF0810466

Name Lise Getoor and David Jacobs, Affiliation University of Maryland, College Park

Problem Statement

While our ability to gather vast amounts of video data is growing at a staggering rate, our ability to effectively store, process, and analyze this video has not kept pace. It is therefore necessary to develop automatic methods for allocating limited resources in video understanding. It particular, it is important to reason about which portions of video require expensive analysis and storage.

We focused on three important video understanding problems. First, we examined low-level vision tasks in which moving objects are separated from the background. Second, we made use of object recognition algorithms for locating and identifying specific objects of interest. Finally, we looked at the problem of identifying activities that are characterized by patterns that occur across space and time, such as repeatedly performing an action. We attack these problems in the context of both single camera video and video from small scale networks. For each of these tasks, we make use of a range of algorithms, both existing and novel, that offer trade-offs in cost and accuracy. We develop inference algorithms that allow us to deploy cheap, noisy algorithms and then reason about which portions of video require more expensive processing.

Summary of Approach

This project aims to make these inferences using new and existing tools from Statistical Relational Learning (SRL). SRL is a recently emerging technology that enables the effective integration of statistical or probabilistic information, with relational or logical domain information, providing the ability to reason collectively about large, complex, interacting domains. Graphical models are used to represent relational information captured in the statistical dependencies between the information available in different portions of video. Then, methods from SRL are used to integrate information in a large video data set and to reason about the label acquisition problem, which tells us which additional information will be most valuable.

We consider both vision algorithms and graphical models that range in from the simple and to the complex. First, we developed systems that can handle background subtraction and object recognition tasks in a single video stream using Hidden Markov Models. This relatively simple graphical model allows us to efficiently perform inference optimally. Next, we developed methods for reasoning about background subtraction and object identification tasks in more complex graphical models with known topology. These can be applied to networks of cameras in which the relationship between images in each camera can be learned offline. Optimal inference is generally intractable in these models, but we will explore the use of approximate algorithms to allow us to control processing in these settings. Finally, we explored the more complex problem in which the topology of the graphical model

must be learned on the fly. This is important for networks with moving cameras, or for inference in very large (eg., gigapixel) cameras, in which a graphical model should be constructed on-the-fly to represent individuals objects of interest and their relationships.

Scientific Barriers

Coping with large amounts of video data presents two fundamental problems. The first concerns computational resources. Modern cameras and networks of cameras can generate such huge volumes of data that it is not possible to apply the most expensive, state-of-the-art algorithms to every frame of video. This is particularly true when we need to evaluate new queries (eg., "Did a blue Ford pickup and a yellow Corrolla ever drive along the same path, stopping at the same parking lot at different times?") on hours or days of video that have been previously collected. In such cases even real-time processing is not adequate, and it is essential to develop methods that control processing, directing it to the portions of video that are most relevant to the task at hand. The second concern is that different portions of large video collections may be related in complex ways. Two distinct portions of video may be related because they show the same object, or related actions. Objects may pass from one camera to another with some time delay, which may vary depending traffic patterns, which vary over time. Our research therefore focuses on developing methods to learn appropriate models of these relationships, and on using these to perform inference and to reason about resource allocation.

Significance

Video surveillance is critical in many military settings. Current systems collect huge amounts of video data, which must be annotated and evaluated by human operators or analysts. Not only is this manual analysis costly, but constrained resources significantly limit the amount of analysis that can be performed. It is also currently very difficult to mine video collections to find unusual or telling patterns (e.g., "Is there a location associated with a significant number of cars that turn around when approaching an unanticipated roadblock?") . Improved methods for automatic video analysis will greatly amplify our ability to process and mine this video.

Summary of Most important Results

- Completed algorithms and experiments for motion detection and object recognition in a single video.
    - Algorithm selects frames to apply expensive detection algorithms. Provably approximates optimal choices very efficiently.
    - Our new algorithms significantly outperform all baseline algorithms, producing more accurate results with much less processing required.

    - Paper published in *NIPS Workshop on Adaptive Sensing, Active Learning and Experimental Design: Theory, Methods and Applications*, 2009.
    - Paper accepted for publication by the *IEEE Transactions on Pattern Analysis and Machine Intelligence.*
- Developed Reflect and Correct algorithm for complex graphical models. Directs human attention to classify a small number of objects that provide the most information about other, related objects.
    - New method outperforms several state-of-the-art algorithms.
    - Paper published in *International Conference on Knowledge Discovery and Data Mining* conference, 2009.
    - Best student paper award.
    - Paper published at the *National Conference on Artificial Intelligence (AAAI NECTAR Track)*, 2010.
    - Paper published in the *ACM Transactions on Knowledge Discovery from Data,* Volume 3, Number 4, page 1-32, November 2009.
- Developed new algorithm for active learning in network data
    - Paper published in *International Conference on Machine Learning*, 2010.
    - Paper on "Query-driven Active Surveying for Collective Classification," with Galileo Mark Namata, Ben London, Lise Getoor, Bert Huang, in *ICML Workshop on Mining and Learning with Graphs*, June 2012.
- Developed new algorithm for enforcing visibility constraints in object recognition from multiple viewpoints.
    - Paper published in *IEEE Conference on Computer Vision and Pattern Recognition, 2009.*
- Developed algorithm for object identification in a camera network.
    - We apply a graphical model to the camera network to allow information to be integrated across space and time.
    - We developed an inference algorithm that can combine low-level image matching and user input to find video frames likely to satisfy a query.
    - We perform *active inference* to determine which frames to ask a human operator to label. We have adapted the Reflect and Correct algorithm so that it can perform this function in camera networks. We show that this algorithm outperforms other approaches.
    - Paper published in the *IEEE Workshop on Person-Oriented Vision, 2011.*
- Developed method for improving object detection by reducing the complexity of latent variable models using group norm regularization. Applied to deformable parts models for detection of people and vehicles.
- Paper submitted to the International Conference on Computer Vision, 2013.
- Three students completed their PhDs at UMD and were supported by this project:
    - Mustafa Bilgic, now assistant professor at TTI
    - Daozheng Chen, now research engineer at Yahoo!
    - Galileo Namata, now research scientist at Verisign

- SRL Tutorials given:
    - Invited Tutorial, Conference on Neural Information Processing (NIPS), December 2012.
    - "Learning Statistical Models from Relational Data," *ACM International Conference on Management of Data (SIGMOD)*, Athens, GR, June, 2011.
    - "Exploiting Statistical & Relational Information on the Web and in Social Media," *Eleventh SIAM International Conference on Data Mining (SDM)*, Phoenix, AZ, April, 2011.
    - "Exploiting Statistical & Relational Information on the Web and in Social Media," *Fourth ACM International Conference on Web Search and Data Mining (WSDM),* Hongkong, CH, February, 2011.
    - "Exploiting Statistical & Relational Information on the Web and in Social Media: Applications, Techniques, and New Frontiers," National Conference on Artificial Intelligence (AAAI), joint with Lily Mihalkova, Atlanta, GA, July, 2010.
- SRL Survey Article
    - *Lifted Graphical Models: A Survey*, Lilyana Mihalkova and Lise Getoor, Machine Learning Journal, 30 pages, accepted subject to minor revisions.

Collaborations and Leveraged Funding

- Lily, NSF funding.
- Object recognition work in collaboration with researchers at the University of Chicago and the Weizmann Institute.
- Object recognition in collaboration with researchers at MIT, TTI and Virginia Tech.

Conclusions

We have demonstrated that inference in graphical models can be used to direct resources to process the most useful pieces of data. We have developed state of the art algorithms for active inference and active learning in network data and have applied them to camera network data.